THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 2, Issue of January 12, pp. 1311–1316, 2001

Printed in U.S.A.

Post-transcriptional Regulation of the *GLI1* Oncogene by the Expression of Alternative 5′ Untranslated Regions*

Received for publication, June 15, 2000, and in revised form, October 10, 2000
Published, JBC Papers in Press, October 13, 2000, DOI 10.1074/jbc.M005191200

Xue-Qing Wang and Joseph A Rothnagel‡

From the Department of Biochemistry and the Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia

The oncogene *GLI1* is involved in the formation of basal cell carcinoma and other tumor types as a result of the aberrant signaling of the Sonic hedgehog-Patched pathway. In this study, we have identified alternative *GLI1* transcripts that differ in their 5′ untranslated regions (UTRs) and are generated by exon skipping. These are denoted α-UTR, β-UTR, and γ-UTR according to the number of noncoding exons possessed (three, two, and one, respectively). The α- and β-UTR forms represent the major *Gli1* transcripts expressed in mouse tissues, whereas the γ-UTR is present at relatively low levels but is markedly induced in mouse skin treated with 12-*O*-tetradecanoylphorbol 13-acetate. Transcripts corresponding to the murine β and γ forms were identified in human tissues, but significantly, only the γ-UTR form was present in basal cell carcinomas and in proliferating cultures of a keratinocyte cell line. Flow cytometry analysis determined that the γ-UTR variant expresses a heterologous reporter gene 14–23-fold higher than the α-UTR and 5–13-fold higher than the β-UTR in a variety of cell types. Because expression of the γ-UTR variant correlates with proliferation, consistent with a role for GLI1 in growth promotion, up-regulation of *GLI1* expression through skipping of 5′ noncoding exons may be an important tumorigenic mechanism.

GLI1 was originally isolated as a highly amplified gene in a malignant glioma (1) and subsequently implicated in the development of other tumor types, including liposarcoma, rhabdomyosarcoma, osteosarcoma, and astrocytoma (2, 3). Later it was shown that *GLI1* encodes a transcription factor that is a downstream nuclear component of the Sonic hedgehog-Patched signaling pathway (4–6). This pathway is evolutionarily conserved and found to operate in a number of tissues during vertebrate development and especially in regions involving mesoderm-ectoderm interactions (7–10). Intercellular signaling by this pathway is initiated when Sonic hedgehog (a secreted protein) binds to Patched (a cell surface transmembrane protein), resulting in the activation of GLI1 in the nucleus and subsequent expression of target genes. Overexpression of *Sonic hedgehog* has been shown to up-regulate *Gli1* in chick limb buds and in the epidermal ectoderm of frog embryos (11, 12), whereas *Gli1* expression is undetectable in *Sonic hedgehog* null mouse embryos (13, 14), confirming that Sonic hedgehog signaling regulates *Gli1* expression.

The discovery of *Patched* mutations in familial and sporadic forms of basal cell carcinoma (BCC),[1] the most common skin cancer, has associated aberrant signaling of the Sonic hedgehog-Patched pathway with the formation of these tumors (15–17). The genetic data are supported by experimental evidence showing that overexpression of *Sonic hedgehog* and other components of this pathway results in the induction of BCCs in transgenic mice and transgenic human skin (18–20). Overexpression of *GLI1* in the epidermis of transgenic animal models produces BCC-like lesions (21, 22) and will transform rodent epithelial cells in cooperation with adenovirus EIA (23), indicating that unregulated expression of *GLI1* is oncogenic. Recent studies have shown that *GLI1* expression is greatly increased in BCCs but not in the surrounding normal tissue, consistent with a central role in tumor formation (21, 24, 25).

In addition to GLI1, two other isoforms have been identified in vertebrates (termed GLI2 and GLI3), each encoded by a separate gene (8, 10). The *GLI* genes are highly expressed during development, and their expression profiles correlate with organogenesis but show only low-level expression in most adult tissues (7, 9, 10). In the skin, *GLI1* expression is readily observed in the epidermal compartment of the developing hair follicle, whereas *GLI2* and *GLI3* transcripts are detected in the surrounding mesenchyme (10, 21, 25). The role of each GLI in mediating the Sonic hedgehog signal is not yet clear, but recent gene ablation studies have shown overlapping roles and indicated some functional redundancy (26, 27). A number of studies have indicated that GLI1 is a transcriptional activator, whereas GLI2 and GLI3 can act as both activators and repressors depending on specific post-translational modifications (28–30). Interestingly, GLI2 and GLI3 are now thought to regulate *GLI1* transcription directly by binding to the *GLI1* promoter (28, 30).

The 5′ untranslation region (UTR) has a large influence on translation and plays a key role in post-transcriptional gene regulation (31–36). The efficiency of translation initiation is largely governed by the composition and structure of the 5′ UTR of the mRNA, which is determined by both its length and its sequence. Extensive secondary structures and one or two small upstream open reading frames (uORFs) within a 5′ UTR can profoundly inhibit protein translation. Most highly expressed mRNAs have relatively short (20–100 nucleotides) 5′ UTRs that lack uORFs and extensive secondary structures (for review, see Ref. 31). In contrast, mRNAs encoding oncopro-

* This work was supported by the Queensland Cancer Fund and the University of Queensland Cancer Research Fund. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by a Wellcome Trust Senior Research Fellowship in Medical Research (Australia). To whom correspondence should be addressed. Tel.: 61-7-3365-4629; Fax: 61-7-3365-4699; E-mail: josephr@biosci.uq.edu.au.

[1] The abbreviations used are: BCC, basal cell carcinoma; GFP, green fluorescent protein; UTR, untranslated region; uORF, upstream open reading frame; RT-PCR, reverse transcription-polymerase chain reaction; RACE, rapid amplification of cDNA ends; mGli, mouse Gli; TPA, 12-*O*-tetradecanoylphorbol 13-acetate.

teins, growth factors, transcription factors, and other regulatory proteins are poorly translated and often have long, highly structured 5′ UTRs with multiple upstream ATGs (32, 35, 36). In this study, we have identified alternative 5′ UTRs of *GLI1* transcripts in mouse and human tissues, which are generated by exon skipping and have marked differences in translation efficiency. Until now, post-transcriptional regulation of *GLI1* has been inferred (37), but a precise mechanism has not been determined. Our results suggest that post-transcriptional regulation of *GLI1* is mediated by exon skipping and show an association of the most efficiently translated 5′ UTR transcript with BCC.

EXPERIMENTAL PROCEDURES

Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR), 5′ Rapid Amplification of cDNA Ends (RACE), and PCR—Skin, brain, heart, kidney, liver, lung, muscle, stomach, spleen, testis, and tongue were obtained from various strains of neonatal, juvenile, and adult mice. Total RNA was isolated from these tissues using TRI Reagent (Molecular Research Center). First-strand cDNA was synthesized from 5 μg of total RNA primed with oligo$(dT)_{16}$ (PerkinElmer) or random hexamers (CLONTECH) using Superscript II reverse transcriptase (Life Technologies) in a total volume of 20 μl. One-tenth by volume of the cDNA was used as the template for subsequent PCR reactions. Primer pairs (listed in Table I) for mouse *GLI1* corresponding to sequences within exon 1 (mGliF1) and exon 2 (mGliR2) and for human *GLI1* corresponding to exon 1 (hGliF1) and 2 (hGliR2) sequences were used. Each PCR reaction was repeated at least three times with different RNA preparations and included negative controls for each set of reactions. For 5′ RACE, we used a cDNA template that was generated using the primer RACE1 and C tailed using terminal transferase according to the manufacturer's protocol (Life Technologies). PCR was performed using the RACE anchor and adapter primers (Life Technologies) and the *GLI1*-specific nested primer RACE2. Mouse genomic DNA was amplified using primers derived from exon 1a (mGliF1a) and exon 2 (mGliR2) sequences. All PCR products were separated on 0.8–2% agarose gels and visualized with ethidium bromide. Separated fragments were purified using QIAEX II (Qiagen) and sequenced directly using the Big Dye termination kit and automated fluorescent sequencing on an ABI-Prism 377 DNA sequencer (PerkinElmer).

12-O-Tetradecanoylphorbol 13-acetate (TPA) Treatment of Mouse Skin—Neonatal and 7-day-old mice (Swiss outbred) were treated with 20–50 μl of 100 μg/ml TPA (Sigma) topically applied to back skin. Mice were killed at different time points (0, 3, 8, and 24 h) after application, and total skin RNA was prepared for RT-PCR as described above.

5′ UTR-Green Fluorescent Protein (GFP) Constructs and Functional Analysis of 5′ UTRs—Each of the three alternative 5′ UTRs of mouse *GLI1* was generated by RT-PCR using primers mGliF1^{Nhe} and mGliQR2^{Age} (Table II) that contain *Nhe*I and *Age*I restriction sites, respectively. PCR products were gel purified and cloned into the pGEM-T Easy vector (Promega). The four ATG codons of the α-UTR were mutated sequentially using the QuikChange site-directed mutageneis kit (Stratagene) and complimentary primers mGliMF1–4 and mGliMR1–4 (Table II). The γ-UTR sequence was multimerized using primers mGliF1^{Bam} and mGliR2^{Bgl} that contain *Bam*HI and *Bgl*II sites, respectively. The amplified fragment was cloned into pGEM-T Easy and sequenced, and the insert was released by digestion with *Bam*HI and *Bgl*II. The purified fragment was ligated in the presence of both restriction enzymes to form similarly oriented concatomers that were then used as templates for PCR amplification using mGLiF1^{Nhe} and mGliR2^{Age}. The products of this reaction were sized on an agarose gel and the band corresponding to four copies of the γ-UTR cloned into the pGEM-T Easy vector. All inserts were verified by sequence analysis, released from the pGEM-T Easy vector by restriction digestion with *Nhe*I and *Age*I, and subcloned into the corresponding sites of the GFP expression vector pEGFP-N1 (CLONTECH).

HaCaT, a human keratinocyte cell line (38), Cos-1 (39), and BHK-21 (40) cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, ampicillin, and streptomycin (Life Technologies). Primary mouse skin fibroblasts were obtained from newborn Swiss mice using an established protocol. Briefly, skin was removed, washed in PBS, and incubated in 2.5% dispase (Life Technologies) for 24 h at 4 °C. The dermis was separated from the epidermis and incubated in 0.2% collagenase (Sigma) at 37 °C for 1 h. Cells were then pelleted, washed in PBS, and cultured in Dulbecco's modified Eagle's medium as described above. Primary fibroblasts were used within the first 2 weeks of culturing.

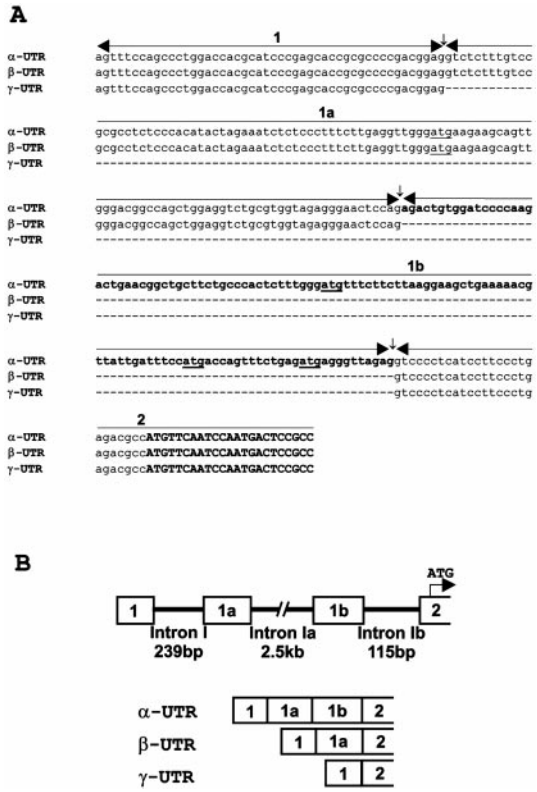

FIG. 1. **Sequence, exon composition, and pre-mRNA structure of the alternative 5′ UTRs of mouse *Gli1*.** *A*, sequence alignment of the three alternative *GLI1* 5′ UTR variants (denoted α-, β-, and γ-UTR) expressed in mouse. The novel 119-base pair sequence of exon 1a is shown in *bold lowercase letters*. The ATG codons denoting the beginning of uORFs are *underlined,* and the main ORF encoding Gli1 is shown in *bold uppercase letters*. The intron-exon boundaries are indicated by *arrows*. *B*, schematic showing the exon composition of the alternative 5′ UTRs and the organization of the pre-mRNA from which they are derived. Exons are denoted by *open boxes,* and introns are denoted by *solid lines* with intron size shown. The translation start site (ATG) of the main ORF is located in exon 2 and indicated by a *bent arrow*.

Transient transfection of GFP constructs was performed using LipofectAMINE Plus reagent (Life Technologies) according to the manufacture's instructions. Cells were seeded on round, glass coverslips in 24-well (fluorescence microscope study) or 6-well (flow cytometry study) plates 24 h before transfection and incubated at 37 °C to a density of ~50–70% confluence. Cultures were washed twice in serum-free media and then incubated in DNA-LipofectAMINE Plus complexes in OPTI-MEM (Life Technologies) for 3 h at 37 °C. Dulbecco's modified Eagle's medium containing serum was then added to the culture. One day later, cells were fixed for microscopy or harvested for flow cytometry analysis. Twenty thousand cells per sample were analyzed on a FACSCalibur (Becton Dickinson) cell sorter, using CELLQuest software (Becton Dickinson).

RESULTS

Identification of Alternative Mouse GLI1 5′ UTRs—To identify the transcriptional start site of the mouse *GLI1* gene, 5′ RACE was performed on total skin RNA from a BALB/c mouse, and the resulting PCR product was sequenced directly. This analysis showed that sequences around the translation start codon and at the beginning of the transcript were identical to the published mouse *GLI1* sequence obtained from F9 cells (41). However, the RACE product also contained an additional 119 bases, not present in the published sequence, located at the splice junction between exons II and III as numbered by Liu *et al.* (41). We named this new 5′ UTR variant α-UTR and the published sequence β-UTR (Fig. 1*A*). To search for other possible splicing variants, RT-PCR was conducted using a forward primer located at exon 1 and a reverse primer that hybridizes

TABLE I
Primer sequences 1

Name	Location[a]	Nucleotide sequence
mGliF1	exon 1, mouse	agtttccagccctggaccacg
mGliR2	exon 2, mouse	ggcgtctcagggaaggatgag
hGliF1	exon 1, human	agactccagccctggaccgcg
hGliR2	exon 2, human	ggcgtctcagaggagggtgtg
RACE1	Exon 4, mouse	gaggtgggaatcctaaag
RACE2	Exon 2/3, mouse	ccagaaagtccttctgttcccatgctgg
mGliF1a	exon 1a, mouse	ctctccctttcttgaggttgg

[a] Exons are numbered according to Figs. 1*B* and 4*B*.

immediately 5′ of the GLI1 ATG codon (Table I). This PCR revealed a much smaller product, which was sequenced directly and found to consist of only one 5′ noncoding exon. This alternative mouse *GLI1* variant corresponds in both size and sequence to the published human transcript obtained from a glioma cell line (7, 41) and was denoted the γ-UTR variant (Fig. 1*A*).

To determine how these alternative transcripts were generated, we examined the genomic organization of this region. Mouse genomic DNA was amplified using the primer pair mGliF1a and mGliR2 (Table I). Sequence analysis of the 2.8-kilobase product identified the novel 119-nucleotide sequence within the α-UTR variant as an authentic exon (which we denoted exon 1b) that is flanked by 2.5 kilobases (intron Ia) and 115 bases (intron Ib) of intervening sequences (Fig. 1*B*). The identification of this additional noncoding exon required a change in the nomenclature used in the earlier study by Lui *et al.* (41), with exon II becoming exon 1a and exon III (which encodes the translation start site) renumbered exon 2 (see Fig. 1*B*).

Expression of Mouse GLI1 5′ UTRs—The expression of the alternative 5′ UTR variants was determined in neonatal, juvenile, and adult tissues from various mouse strains by RT-PCR (Fig. 2). This analysis revealed that the UTR variants had no particular tissue-specific expression pattern but did show marked strain-specific differences. In all tissues examined the larger UTR forms (α- and β-UTR) predominate, whereas the γ-UTR appears as a minor amplified product (Fig. 2*A*). In some strains, such as BALB/c, DBA, and C57BL/6, the α-UTR variant was the major form, whereas in CD1 and SV129 strains, the β-UTR form was the predominant transcript (Fig. 2*B*). In Swiss outbred mice, expression of the α- and β-UTR was heterogeneous, with some animals expressing both forms and others expressing only one (data not shown). In a given individual, the expression profile of the two larger transcripts was identical in all tissues examined irrespective of the strain used (Fig. 2*A*). We also followed the expression of the UTR variants in postnatal skin development and found that the apparent levels of all *GLI1* transcripts were reduced with increasing age and that the γ-UTR transcript was not detected at all in adult skin (data not shown).

To evaluate whether the expression of these 5′ UTR variants correlated with proliferative status, newborn and 7-day-old animals were treated with TPA topically applied to back skin. These experiments revealed that expression of the larger UTR transcripts was maximally reduced at 3 h after application, whereas expression of the γ form was increased (Fig. 3). Reduced expression of the α- and β-UTRs was still evident, albeit not as marked, 24 h after application (data not shown). Because acute TPA treatment results in increased mitotic activity of basal layer keratinocytes (42), these data indicate an association of the γ-UTR transcript with proliferation and the α- and β-UTRs with differentiation.

Identification of Alternative Human GLI1 5′ UTRs—We next searched for alternative human *GLI1* transcripts in newborn foreskin by RT-PCR with primers derived from exons 1 and 2 (Table I) of the published sequence (7). Two PCR products were

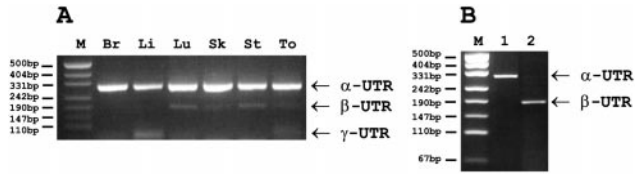

FIG. 2. **The expression of alternative *GLI1* 5′ UTRs is not tissue specific but does show strain variation.** RT-PCR was performed on mRNA isolated from brain (*Br*), liver (*Li*), Lung (*Lu*), skin (*Sk*), stomach (*St*), and tongue (*To*) tissues of a postnatal mouse (*A*). A signal for the β- and γ-UTR variants is also present in all tissues examined but is barely discernible. *B*, in BALB/c, DBA, and C57Bl/6 strains, the α-UTR variant is the major transcript (BALB/c is shown in *lane 1*), whereas the β-UTR variant is the dominant transcript in CD-1 and SV129 strains (CD-1 is shown in *lane 2*). *M*, DNA marker.

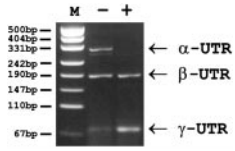

FIG. 3. **Expression of the *GLI1* 5′ UTR variants is altered by TPA treatment.** The expression of the γ-UTR transcript is increased, whereas expression of the α-UTR is reduced, in TPA-treated skin (+) relative to control skin (−). *M*, DNA size marker.

generated and sequenced, with the smaller fragment corresponding to the published sequence (7) and the larger containing an additional 144 bases located at the splice junction of exons 1 and 2 (Fig. 4*A*). The larger transcript was termed β-UTR, and the smaller sequence was termed the γ-UTR variant (Fig. 4, *A* and *B*). Notably, the novel 144-nucleotide sequence found within the human β-UTR has significant homology with mouse exon 1a but is 30 bases larger and contains two additional ATG codons (Fig. 4*C*).

The β- and γ-UTRs were also present in a brain sample, and their identity was confirmed by sequence analysis (Fig. 5). The expression of the β- and γ-UTRs was further examined in HaCaT cells and seven BCC samples. We found that the γ transcript was present in proliferating cultures of HaCaT cells and all BBC samples, but in contrast to foreskin keratinocytes and brain tissue, we were unable to amplify the β transcript from these mRNAs (Fig. 5). Therefore, the γ-UTR transcript may represent the major variant expressed by proliferating cells in human tissues as well.

Functional Analysis of the GLI1 5′ UTRs—The 5′ UTR is known to regulate gene expression by influencing the efficiency of translation. An examination of the *GLI1* 5′ UTR variants revealed three small uORFs in mouse α-UTR, two in human β-UTR, one that overlaps the GLI1 ORF in the mouse β-UTR, and none in the γ-UTRs (Table III). The secondary structure for each 5′ UTR was analyzed using the RNA folding prediction program MFOLD (43). This program predicted extensive secondary structures in the longer UTRs, with calculated free energy values of −72.7 to −94.5 kcal/mol for the mouse α-UTR, −65 kcal/mol for the human β-UTR, and −55 kcal/mol for the mouse β-UTR. The human and mouse γ-UTRs were predicted not to form stable secondary structures and to have free energy values of −12 kcal/mol (Table III). The presence of uORFs and stable secondary structures in the larger UTRs suggests that they will be less efficiently translated than the γ variant, which lacks these features.

To test the above prediction *in vivo*, the three mouse *GLI1* 5′ UTR fragments were cloned upstream of a GFP reporter gene (Fig. 6*A*), and the constructs were transiently transfected into HaCaT, Cos-1, and BHK-21 mammalian cell lines and primary mouse skin fibroblasts. The difference in GFP fluorescence produced by these constructs was striking (Fig. 6*B*), with the

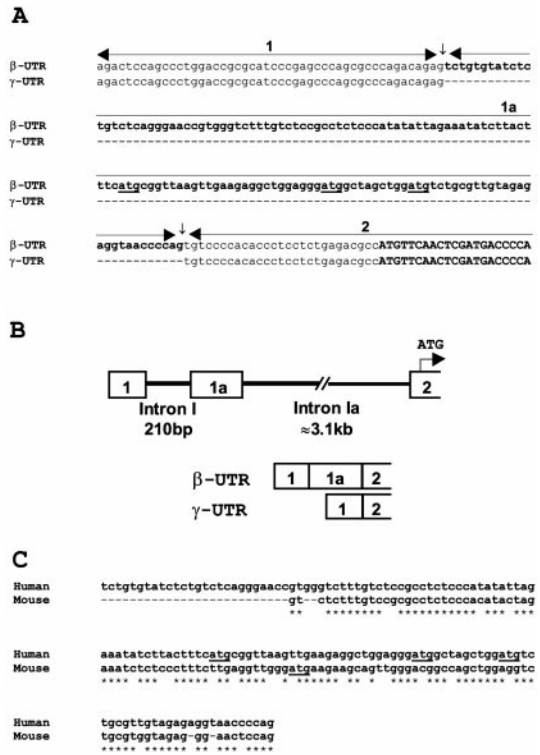

FIG. 4. **Sequence, exon composition, and pre-mRNA structure of alternative human *GLI1* 5′ UTRs.** *A*, sequence alignment of the alternative *GLI1* 5′ UTR variants identified in human tissues (denoted β- and γ-UTR). The novel 144-base pair sequence of exon 1a is shown in *bold lowercase letters*. The ATG codons denoting the beginning of uORFs are *underlined,* and the main ORF encoding GLI1 is shown in *bold uppercase letters*. The intron-exon boundaries are indicated by *arrows*. *B*, schematic showing the exon composition of the alternative 5′ UTRs and the organization of the pre-mRNA from which they are derived. Exons are denoted by *open boxes,* and introns are denoted by *solid lines* with intron size shown. The translation start site (ATG) of the main ORF is located in exon 2 and indicated by a *bent arrow*. *C,* comparison of human exon 1a sequences with mouse exon 1a. ∗, identical bases.

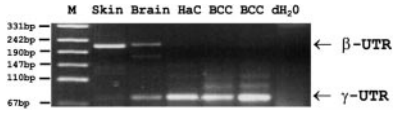

FIG. 5. **Detection of human *GLI1* 5′ UTR variants by RT-PCR.** RT-PCR results from skin, brain, the HaCaT human keratinocyte cell line (*HaC*), two BCC biopsies, and a no-RNA control (dH_2O) are shown. *M*, DNA size marker.

α-UTR construct producing the lowest fluorescence levels, the β-UTR construct producing intermediate levels, and the γ variant producing the highest levels (even higher than the GFP vector control). Importantly, there was no apparent difference in the transfection efficiency of these constructs for a given cell type, showing that the increase in the number of brightly fluorescing cells transfected with the γ form is caused by an increase in GFP production.

To quantify these results, cells were also subjected to flow cytometry analysis (Fig. 7*A*). The advantages of this technique over enzymic reporter assays are that GFP levels are determined in individual cells using a large number of cells (20,000 cells/construct), and untransfected cells are discarded from the calculations, thus removing any bias attributable to transfection efficiencies. This analysis revealed that the α- and β-UTRs expressed the reporter 60–90% lower than the GFP vector control, whereas the γ-UTR produced 2–3-fold enhancement of expression (Fig. 7*B*). Comparison of expression levels between the three variants reveals a 14–23-fold increase in GFP production by the mouse γ-UTR construct over the α-UTR (γ/α) and a 5–13-fold increase over the β-UTR (γ/β; Fig. 7*C*). The greatest differences in GFP intensities were seen in HaCaT cells. These data show that the γ-UTR facilitates expression of a heterologous protein, whereas the α- and β-UTRs significantly suppress protein production.

TABLE II
Primer sequences 2

Names	Nucleotide sequence[a]
mGliF1^{Nhe}	**gctagc**agtttccagccctggaccacg
mGliR2^{Age}	**accggt**ggcgtctcagggaaggatgag
mGliF1^{Bam}	**ggatcc**agtttccagccctggaccacg
mGliR2^{Bgl}	**agatct**ggcgtctcagggaaggatgag
mGliMF1	tcttgaggttgggttgaagaagcagtt
mGliMR1	aactgcttcttcaacccaacctcaaga
mGliMF2	cccactctttgggttgtttcttcttaa
mGliMR2	ttaagaagaaacaacccaaagagtggg
mGliMF3	gttattgatttccttgaccagtttctg
mGliMR3	cagaaactggtcaaggaaatcaataac
mGliMF4	accagtttctgagttgagggttagagg
mGliMR4	cctctaaccctcaactcagaaactggt

[a] Primers mGliF1^{Nhe} and mGliR2^{Age} and mGliF1^{Bam} and mGliR2^{Bgl} are identical to mGliF1 and mGliR2 (Table I) but include restriction sites for *Nhe*I, *Age*I, *Bam*HI, and *Bgl*II, respectively (bold). The point mutations (A → T) introduced by primers mGliMF1–4 and mGliMR1–4 are *underlined*.

To determine whether this suppression is caused by the presence of uORFs or some other property of the longer UTRs, such as increased secondary structure, we made two additional constructs. The uORFs of the α-UTR variant were removed by mutating all four upstream ATG codons to TTG. Significantly, the mutant α-UTR construct expressed the reporter at levels that approached that of the γ-UTR construct (Fig. 8). To ascertain whether the length of the UTR sequence could influence expression levels, we produced a γ-UTR multimer containing four copies of this sequence in the same orientation. The γ-UTR multimer, which contained 314 base pairs compared with the 307 base pairs of the α- variant, produced GFP levels that were only marginally lower than a single copy of the γ-UTR (Fig. 8). Taken together, these data show that the uORFs of the Gli1 UTRs play a major role in the suppression of protein production.

DISCUSSION

The expression of alternative mRNAs from a single gene is an important mechanism for gene regulation and for the generation of functionally different proteins (44, 45). In this study, we have identified alternative *GLI1* transcripts that differ in their 5′ UTRs and are generated by exon skipping. The longer transcripts (α and β in mouse and β in human) were present in all normal tissues examined, whereas the shortest transcript (γ-UTR) was present in neonatal tissues but rarely detected in adult tissues. Significantly, the γ form is the only variant found in BCCs and in a proliferating human keratinocyte cell line. Moreover, expression of the γ form could be induced by TPA treatment of mouse skin. We have found that the *GLI1* 5′ UTRs determine the expression levels of a heterologous coding sequence in transfected cells, with the γ form associated with the highest levels of expression and the α- and β-UTRs associated with significantly lower levels. These findings suggest that GLI1 levels may be regulated through the use of alternative 5′ UTRs and predict that an unregulated increase in the γ transcript over the α and β forms may be tumorigenic.

We initially examined several tissues from various strains of mice, both neonatal and adult, for evidence of tissue-specific expression of the *GLI1* alternative transcripts. We found that expression of these transcripts did not show tissue specificity, but the levels of the α and β transcripts (relative to each other) did vary between strains. Because BALB/c mice predominantly

express the α form and CD1 mice predominantly express the β form, we sequenced the intron-exon boundaries of this region in these strains and found them to be identical (data not shown), suggesting that the expression differences between strains was not caused by polymorphic variation at these sites. Whether these strain differences were caused by polymorphic differences in intronic *cis*-acting elements or by differences in trans-acting factors remains to be determined. We did not observe any transcripts equivalent to the mouse α-UTR in the human tissues examined, and we were unable to find sequences equivalent to mouse exon 1b in the human gene, suggesting that humans do not have the capacity to express an α-UTR variant, although the presence of three upstream ATG codons in exon 1a of the human message compared with just one in mouse, together with the increased stability of the human β-UTR (Table III), appears to compensate for the lack of exon 1b sequences in the human sequence. However, we cannot exclude the possibility that expression of the α-UTR variant may be idiosyncratic in human as it is in mice because of the small sample size that was available in this study.

The expression of the 5′ UTR variants did show an apparent correlation with proliferative status. The longest transcripts are expressed in normal tissues but not in BCCs or in a human keratinocyte cell line, whereas the shortest transcript is the only variant present in BCCs and in HaCaT cells. In addition, expression of the γ form was induced by TPA treatment of mouse skin, whereas the levels of the longer transcripts were concomitantly reduced. These observations suggest an association of the shortest 5′ UTR with actively proliferating keratinocytes and the α and β transcripts with quiescent cells. The association of the most actively translated transcript with pro-

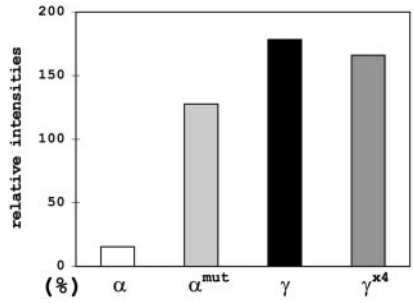

FIG. 8. **Flow cytometry analysis of mutant 5′ UTR-GFP constructs in transfected cells.** A graphical representation of GFP intensities of wild-type α-UTR (α), mutant α-UTR (α^{mut}), wild-type γ-UTR (γ), and the 4-mer γ-UTR (γ^{x4}) constructs relative to the empty GFP vector in Cos-1 cells is shown. The values shown are the averages of three independent transfection and cell-sorting experiments.

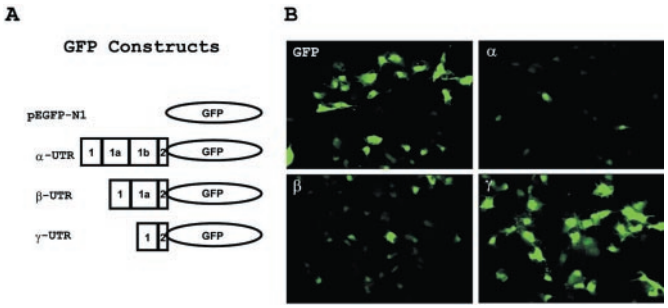

FIG. 6. **The *Gli1* 5′ UTR variants differentially express a GFP reporter construct in transfected cells.** *A*, diagram showing the GFP constructs used in the transfection studies. The mouse *Gli1* 5′ UTR sequences were cloned upstream of the GFP ORF as indicated. *pEGFP-N1,* parent GFP construct. *B*, GFP fluorescence observed in Cos-1 cells transfected with the GFP expression vector alone (*GFP*) and α-UTR-GFP (α), β-UTR-GFP (β), and γ-UTR-GFP (γ) constructs. Transfection efficiency was determined by counting GFP-expressing cells, which revealed that each construct was transfected with the same efficiency (within experimental limits).

TABLE III
Theoretical analysis of GLI1 5′ UTRs

5′ UTRs	uATGs[a]	uORFs[b]	Size of uORFs (no. of codons)	Free energy[c]
				Kcal/mol
Mouse α-UTR	4	3	26,4,10	−73 to −95
Mouse β-UTR	1	1	38	~−55
Mouse γ-UTR	0	0	—	~−12
Human β-UTR	3	2	4, 8	~−65
Human γ-UTR	0	0	—	~−12

[a] Number of ATG codons present upstream of the authentic translation start site.

[b] Number of ORFs found upstream of the authentic translation start site. Note that only the uORF in mouse β-UTR overlaps with the main *GLI1* ORF.

[c] The mean free energy values are indicated for each UTR except for mouse α-UTR, for which the lowest and highest values are given. These were calculated using MFOLD (43).

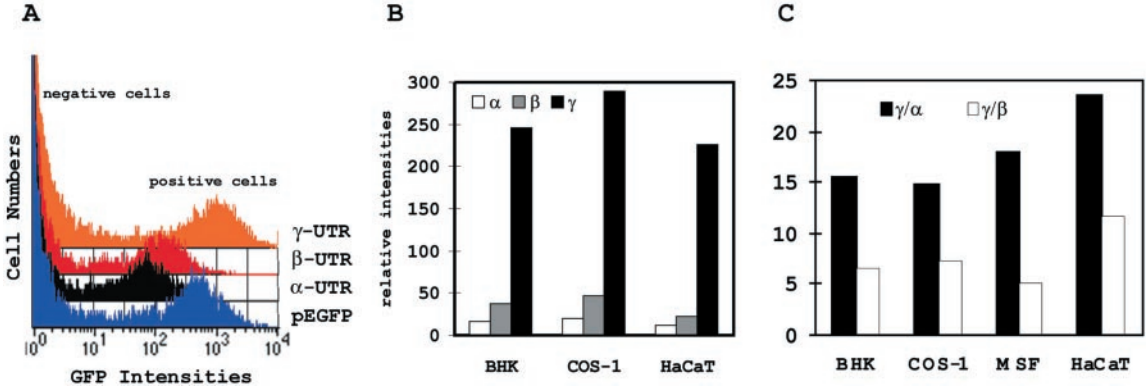

FIG. 7. **Flow cytometry analysis of 5′ UTR-GFP constructs in transfected cells.** *A*, fluorescence intensity histogram compiled from the analysis of 20,000 cells per sample for each construct (see Fig. 6) in transfected Cos-1 cells. The peak closest to the *vertical axis* is attributable to untransfected cells (negative cells), which show low fluorescent intensities because of autofluorescence. The second peak represents fluorescence from transfected cells that express GFP. The four histograms representing each GFP construct were merged to allow direct comparison of fluorescence intensities. This analysis revealed that the γ-UTR construct produced the highest GFP intensities (mean value, 1.045×10^3) and the α-UTR produced the lowest (mean value, 6.2×10^1). *B*, graphical representation of GFP intensities of the 5′ UTR variants relative to the empty GFP vector (defined as 100) in transfected BHK-21 (*BHK*), Cos-1, and HaCaT cells. *C*, graphical representation of the ratios of GFP intensities among 5′ UTR variants in the cell lines indicated. The γ-UTR/α-UTR (■ γ/α) and γ-UTR/β-UTR (□ γ/β) ratios are shown for constructs transfected into BHK-21 cells (*BHK*), Cos-1 cells, primary mouse skin fibroblasts (*MSF*), and HaCaT cells. The largest differences in GFP fluorescence between the γ-UTR and the α- and β-UTR variants were observed in HaCaT cells. The values shown are the averages of three independent transfection and cell sorting experiments.

liferation fits neatly with the observation of increased GLI1 expression in BCCs and with experimental *in vivo* data showing overexpression of *GLI1* resulting in BCC-like lesions (21, 22, 24, 25). The concomitant reduction in α and β transcripts with increased levels of the γ form after TPA treatment is consistent with a post-transcriptional modification of a single pre-mRNA species. Although the precise mechanism is unknown, the activity of transacting factors that control alternative splicing decisions have been shown to be modulated by mitogens such as TPA (46–49). It can also be reasoned that alternative splicing of the *GLI1* message may be regulated by Sonic hedgehog signaling. We are currently testing the possibility that unregulated signaling of the Sonic hedgehog-Patched pathway (as occurs in BCCs) results in the generation of increased levels of the γ form through changes in the composition of splicing factors bound to the nascent *GLI1* pre-mRNA. It is also conceivable that tumor-causing mutations could occur in the splice sites of upstream noncoding exons that would enhance the formation of the γ form. Alternatively, mutations that alter the regulation of factors involved in splice site selection may also contribute to tumor formation. In addition, GLI1 has been shown to bind RNA directly at a conserved motif that is distinct from the GLI DNA binding *cis*-element and to regulate translation of the bound mRNA (37, 50). Because these putative GLI binding motifs are present in the pre-mRNA sequence of *GLI1*, an intriguing possibility is that GLI1 itself or a related protein may determine exon selection. Two recent reports have indicated that GLI2 and GLI3 transcriptionally regulate *GLI1* by binding directly to promoter sequences, but the same data do not preclude up-regulation of *GLI1* via a post-transcriptional mechanism by these factors as well (28, 30).

Our results suggest that the α- and β-UTR variants contain potent translation inhibitors and are in agreement with other studies showing that 5′ UTRs with uORFs and stable secondary structures are features of poorly translated mRNAs (31–36). Furthermore, we have shown that the uORFs present in the α- and β-UTRs play an important role in mediating this inhibition. The increased GFP levels observed for the γ form may be wholly attributable to the lack of uORFs, because increasing the length of the γ-UTR variant 4-fold only marginally decreased GFP levels. However, it is also possible that the γ-UTR transcript may be more stable than the longer variants and consequently is able to produce more product per transcript than the α- and β-UTR variants. Several studies have shown that impaired translation can initiate mRNA decay and that both stem-loop structures and uORFs within 5′ UTRs can modulate mRNA stability (51). Regardless of whether these sequences modulate translation efficiency or mRNA stability (or both), the retention or skipping of 5′ noncoding exons generates alternative UTRs that differentially express a heterologous reporter, and it is assumed of GLI1 as well. In conclusion, this study suggests that alternative splicing is an important mechanism in the regulation of this gene, which may in part contribute to the up-regulation of *GLI1* observed in BCCs.

Acknowledgments—We thank Dr. Norbert Fusenig for the HaCaT cells, Dr. Peter Dodd for the human brain mRNA sample, and Dr. Anthony Dicker for human foreskin biopsies and dissected BCC samples. We thank Betsy Hung and Lexie Friend for assistance with cell culture.

REFERENCES

1. Kinzler, K. W., Bigner, S. H., Bigner, D. D., Trent, J. M., Law, M. L., O'Brien, S. J., Wong, A. J., and Vogelstein, B. (1987) *Science* **236,** 70–73
2. Roberts, W. M., Douglass, E. C., Peiper, S. C., Houghton, P. J., and Look, A. T. (1989) *Cancer Res.* **49,** 5407–5413
3. Stein, U., Eder, C., Karsten, U., Haensch, W., Walther, W., and Schlag, P. M. (1999) *Cancer Res.* **59,** 1890–1895
4. Ingham, P. W. (1998) *EMBO J.* **17,** 3505–3511
5. Johnson, R. L., and Scott, M. P. (1998) *Curr. Opin. Genet. Dev.* **8,** 450–456
6. Ruiz i Altaba, A. (1999) *Nat. Cell Biol.* **1,** 147–148
7. Kinzler, K. W., Bigner, S. H., Ruppert, J. M., and Vogelstein, B. (1988) *Nature* **332,** 371–374
8. Ruppert, J. M., Kinzler, K. W., Wong, A. J., Bigner, S. H., Kao, F. T., Law, M. L., Seuanez, H. N., O'Brien, S. J., and Vogelstein, B. (1988) *Mol. Cell. Biol.* **8,** 3104–3113
9. Walterhouse, D., Ahmed, M., Slusarski, D., Kalamaras, J., Boucher, D., Holmgren, R., and Iannaccone, P. (1993) *Dev. Dyn.* **196,** 91–102
10. Hui, C. C., Slusarski, D., Platt, K. A., Holmgren, R., and Joyner, A. L. (1994) *Dev. Biol.* **162,** 402–413
11. Marigo, V., Johnson, R. L., Vortkamp, A., and Tabin, C. J. (1996) *Dev. Biol.* **180,** 273–283
12. Lee, J., Platt, K. A., Censullo, P., and Ruiz i Altaba, A. (1997) *Development* **124,** 2537–2552
13. St-Jacques, B., Dassule, H. R., Karavanova, I., Botchkarev, V. A., Li, J., Danielian, P. S., McMahon, J. A., Lewis, P. M., Paus, R., and McMahon, A. P. (1998) *Curr. Biol.* **8,** 1058–1068
14. Chiang, C., Swan, R. Z., Grachtchouk, M., Bolinger, M., Litingtung, Y., Robertson, E. K., Cooper, M. K., Gaffield, W., Westphal, H., Beachy, P. A., and Dlugosz, A. A. (1999) *Dev. Biol.* **205,** 1–9
15. Gailani, M. R., Stahle-Backdahl, M., Leffell, D. J., Glynn, M., Zaphiropoulos, P. G., Pressman, C., Unden, A. B., Dean, M., Brash, D. E., Bale, A. E., and Toftgard, R. (1996) *Nat. Genet.* **14,** 78–81
16. Hahn, H., Wicking, C., Zaphiropoulous, P. G., Gailani, M. R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Unden, A. B., Gillies, S., Negus, K., Smyth, I., Pressman, C., Leffell, D. J., Gerrard, B., Goldstein, A. M., Dean, M., Toftgard, R., Chenevix-Trench, G., Wainwright, B., and Bale, A. E. (1996) *Cell* **85,** 841–851
17. Johnson, R. L., Rothman, A. L., Xie, J., Goodrich, L. V., Bare, J. W., Bonifas, J. M., Quinn, A. G., Myers, R. M., Cox, D. R., Epstein, E. H., Jr., and Scott, M. P. (1996) *Science* **272,** 1668–1671
18. Fan, H., Oro, A. E., Scott, M. P., and Khavari, P. A. (1997) *Nat. Med.* **3,** 788–792
19. Oro, A. E., Higgins, K. M., Hu, Z., Bonifas, J. M., Epstein, E. H., Jr., and Scott, M. P. (1997) *Science* **276,** 817–821
20. Xie, J. M., Murone, M., Luoh, S. M., Ryan, A., Gu, Q. M., Zhang, C. H., Bonifas, J. M., Lam, C. W., Hynes, M., Goddard, A., Rosenthal, A., Epstein, E. H., Jr., and de Sauvage, F. J. (1998) *Nature* **391,** 90–92
21. Dahmane, N., Lee, J., Robins, P., Heller, P., and Ruiz i Altaba, A. (1997) *Nature* **389,** 876–881
22. Nilsson, M., Unden, A. B., Krause, D., Malmqwist, U., Raza, K., Zaphiropoulos, P. G., and Toftgard, R. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 3438–3443
23. Ruppert, J. M., Vogelstein, B., and Kinzler, K. W. (1991) *Mol. Cell. Biol.* **11,** 1724–1728
24. Reifenberger, J., Wolter, M., Weber, R. G., Megahed, M., Ruzicka, T., Lichter, P., and Reifenberger, G. (1998) *Cancer Res.* **58,** 1798–1803
25. Ghali, L., Wong, S. T., Green, J., Tidman, N., and Quinn, A. G. (1999) *J. Invest. Dermatol.* **113,** 595–599
26. Motoyama, J., Liu, J., Mo, R., Ding, Q., Post, M., and Hui, C. C. (1998) *Nat. Genet.* **20,** 54–57
27. Park, H. L., Bai, C., Platt, K. A., Maise, M. P., Beeghly, A., Hui, C. C., Nakashima, M., and Joyner, A. L. (2000) *Development* **127,** 1593–1605
28. Dai, P., Akimaru, H., Tanaka, Y., Maekawa, T., Nakafuku, M., and Ishii, S. (1999) *J. Biol. Chem.* **274,** 8143–8152
29. Ruiz I Altaba, A. (1999) *Development* **126,** 3205–3216
30. Sasaki, H., Nishizaki, Y., Hui, C. C., Nakafuku, M., and Kondoh, H. (1999) *Development* **126,** 3915–3924
31. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148
32. Kozak, M. (1991) *J. Cell Biol.* **115,** 887–903
33. Sonenberg, N. (1994) *Curr. Opin. Genet. Dev.* **4,** 310–315
34. Kozak, M. (1996) *Mamm. Genome* **7,** 563–574
35. van der Velden, A. W., and Thomas, A. A. (1999) *Int. J. Biochem. Cell Biol.* **31,** 87–106
36. Willis, A. E. (1999) *Int. J. Biochem. Cell Biol.* **31,** 73–86
37. Graves, L. E., Segal, S., and Goodwin, E. B. (1999) *Nature* **399,** 802–805
38. Boukamp, P., Petrussevska, R. T., Breitkreutz, D., Hornung, J., Markham, A., and Fusenig, N. E. (1988) *J. Cell Biol.* **106,** 761–771
39. Gluzman, Y. (1981) *Cell* **23,** 175–182
40. Ferrari, M. (1962) *Virology* **16,** 147–151
41. Liu, C. Z., Yang, J. T., Yoon, J. W., Villavicencio, E., Pfendler, K., Walterhouse, D., and Iannaccone, P. (1998) *Gene (Amst.)* **209,** 1–11
42. Heyden, A., Lutzow-Holm, C., Clausen, O. P., Thrane, E. V., Brandtzaeg, P., Roop, D. R., Yuspa, S. H., and Huitfeldt, H. S. (1994) *Differentiation* **57,** 187–193
43. Zuker, M. (1989) *Science* **244,** 48–52
44. Ayoubi, T. A., and Van De Ven, W. J. (1996) *FASEB J.* **10,** 453–460
45. Lopez, A. J. (1998) *Annu. Rev. Genet.* **32,** 279–305
46. Darville, M. I., and Rousseau, G. G. (1997) *Nucleic Acids Res.* **25,** 2759–2765
47. Du, K., Leu, J. I., Peng, Y., and Taub, R. (1998) *J. Biol. Chem.* **273,** 35208–35215
48. Levanon, D., Bernstein, Y., Negreanu, V., Ghozi, M. C., Bar-Am, I., Aloya, R., Goldenberg, D., Lotem, J., and Groner, Y. (1996) *DNA Cell Biol.* **15,** 175–185
49. Screaton, G. R., Caceres, J. F., Mayeda, A., Bell, M. V., Plebanski, M., Jackson, D. G., Bell, J. I., and Krainer, A. R. (1995) *EMBO J.* **14,** 4336–4349
50. Jan, E., Yoon, J. W., Walterhouse, D., Iannaccone, P., and Goodwin, E. B. (1997) *EMBO J.* **16,** 6301–6313
51. Jacobson, A., and Peltz, S. W. (1996) *Annu. Rev. Biochem.* **65,** 693–739

Post-transcriptional Regulation of the *GLI1* Oncogene by the Expression of Alternative 5′ Untranslated Regions

Xue-Qing Wang and Joseph A Rothnagel

J. Biol. Chem. 2001, 276:1311-1316.
doi: 10.1074/jbc.M005191200 originally published online October 13, 2000